
 RELEASE 3.1A JOHN F. COLLINS, BIOCOMPUTING RESEARCH UNIT.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Mparch_pp protein - protein database search, using Smith-Waterman algorithm
 R : Thu Apr 27 09:32:02 2000; Masspar time 41.60 seconds
 Tabular output not generated. 486.838 Million cell updates/sec

Title: >US-09-421-213-2
 Description: (1-855) from US09421213B.pep
 Perfect Score: 6382
 Sequence: 1 MGSDDRRKGGGGKPDGAGL.....PGVYTRLPFLFRDWIKENTGV 855

Scoring table: PAM 150
 Gap 11

Searched: 188963 seqs, 23686106 residues

Post-Processing: Minimum Match 0%
 Listing first 45 summaries

Database: a-geneseq36
 1:geneseqp

Statistics: Mean 38.523; Variance 166.585; scale 0.231

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

P	Score	Query Match	Length	ID	Description	Pred. No.
1	1808	28.3	241	1 W22987	Human serine protease	8.25e-161
2	768	12.0	798	1 R57283	Bovine enterokinase.	2.72e-59
3	677	10.6	418	1 R89435	Trypsin-like enzyme.	1.35e-50
4	663	10.4	232	1 R89430	Trypsin-like enzyme.	3.01e-49
5	618	9.7	416	1 W96812	A mouse serine protease	5.72e-45
6	615	9.6	356	1 W46917	Amino acid sequence of	1.10e-44
7	604	9.4	297	1 W77304	Amino acid sequence of	2.91e-43
8	578	9.1	812	1 R83959	Complete mouse plasmin	3.52e-41
9	578	9.1	812	1 W94036	Murine plasminogen.	3.52e-41
10	578	9.1	812	1 W07585	Murine plasminogen.	3.52e-41
11	567	8.9	271	1 W77302	Amino acid sequence of	3.85e-40
12	567	8.9	317	1 Y13391	Amino acid sequence of	3.85e-40
13	566	8.9	437	1 W51457	Human plasminogen	4.79e-40
14	566	8.9	437	1 W51457	Human plasminogen	4.79e-40
15	566	8.9	437	1 W51457	Human plasminogen	4.79e-40
16	567	8.9	791	1 W34285	Human plasminogen.	3.85e-40
17	566	8.9	810	1 R06065	Human plasminogen.	4.79e-40
18	566	8.9	810	1 W31169	Plasminogen protein fo	4.79e-40
19	566	8.9	810	1 R34428	Sequence encoded by a	5.95e-40
20	565	8.9	811	1 R12933	Plasminogen muten X1	5.95e-40
21	565	8.9	812	1 R12934	Plasminogen muten X2	5.95e-40
22	565	8.9	813	1 R12935	Plasminogen muten X3	5.95e-40
23	568	8.9	816	1 R12946	Plasminogen muten T19	3.10e-40

24	561	8.8	233	1 W22986	Human serine protease	1.42e-39
25	560	8.8	461	1 W40284	Human factor IX protea	1.76e-39
26	562	8.8	810	1 R12938	Plasminogen muten T1	1.14e-39
27	562	8.8	811	1 R12943	Plasminogen muten T1	1.14e-39
28	562	8.8	811	1 R12944	Plasminogen muten T1	1.14e-39
29	562	8.8	811	1 R12945	Plasminogen muten T1	1.14e-39
30	562	8.8	811	1 R12946	Plasminogen muten T1	1.14e-39
31	562	8.8	813	1 R12948	Plasminogen muten T2	1.14e-39
32	563	8.8	813	1 R12948	Plasminogen muten T2	1.14e-39
33	561	8.8	813	1 R12937	Plasminogen muten X5	9.19e-40
34	566	8.7	415	1 Y03203	Plasminogen muten X6	1.42e-39
35	563	8.7	415	1 R35761	Amino acid sequence of	4.21e-39
36	565	8.7	454	1 R67710	Factor IX (IX).	8.07e-39
37	555	8.7	456	1 P40178	Human factor-IX.	5.23e-39
38	555	8.7	461	1 R05393	Part of the sequence o	5.23e-39
39	555	8.7	461	1 P50311	Mutant human factor IX	5.23e-39
40	555	8.7	461	1 P50312	Sequence of human fac	5.23e-39
41	557	8.7	810	1 R13221	Human plasminogen vari	3.39e-39
42	556	8.7	810	1 R13220	R5615 human plasminoge	4.21e-39
43	556	8.7	810	1 R12406	R5612 human plasminoge	4.21e-39
44	556	8.7	810	1 R12319	R5616 human plasminoge	4.21e-39
45	558	8.7	816	1 R12942	Plasminogen muten T8	2.72e-39

ALIGNMENTS

RESULT	1	ALIGNMENTS
ID	W22987 standard; Protein: 241 AA.	
AC	W22987.	
DT	08-OCT-1997 (first entry)	
DE	Human serine protease 67 (SP67).	
KW	Human; colon carcinoma; COLO 201; cell line; serine protease; SP67;	
KV	screening; inhibitor; treatment; disease.	
OS	Homo sapiens.	
PN	J09149/90-A.	
PD	10-JUN-1997.	
PR	24-JUL-1996; 212196.	
PR	29-SEP-1995; JP-275105.	
PA	(SUNR) SUNITORY LTD.	
DR	WPI:97-357902/33.	
DR	N-PSDB: T79128.	
PT	Human colon carcinoma derived serine protease(s) SP59, SP60 and SP67	
PT	- useful to screen for specific inhibitors, e.g. to search for, or	
PT	study agent for treatment of various diseases	
PS	Claim 1; Pages 12-13; 16pp; Japanese.	
CC	The present sequence is the human colon carcinoma COLO 201	
CC	cell line derived serine protease 67 (SP67), which can be used to	
CC	screen for specific inhibitors, e.g. to search for, or study an	
CC	agent for the treatment of various diseases.	
SC	Sequence 241 AA;	

Query Match	28.3%	Score 1808; DB 1; Length 241;
Best Local Similarity	99.6%;	Pred. No. 8.25e-161;
Matches 240; Conservative	0; Mismatches 1; Indels 0; Gaps 0	

DB	1	VVGGDDADESMPQVSHALGOGHICASLISPMWLSAACHYIDDGFRSDPTQTV 60
OY	615	VVGGDDADESMPQVSHALGOGHICASLISPMWLSAACHYIDDGFRSDPTQTV 674
OY	615	VVGGDDADESMPQVSHALGOGHICASLISPMWLSAACHYIDDGFRSDPTQTV 674
DB	61	FLGLHDQSGASAPGOERLKRITSHPFNFTDYDIALLEKPAVSSWVRICPD 120
OY	675	FLGLHDQSGASAPGOERLKRITSHPFNFTDYDIALLEKPAVSSWVRICPD 734
OY	735	ASHFFPACAIWVGWGTGTGTCGALLTGKEIRVINOTTCENLLPOQIPRACVGL 794
DB	121	ASHFFPACAIWVGWGTGTGTCGALLTGKEIRVINOTTCENLLPOQIPRACVGL 180
OY	735	ASHFFPACAIWVGWGTGTGTCGALLTGKEIRVINOTTCENLLPOQIPRACVGL 794
DB	181	SGVDSCGSGGPGPLSSVEADGRIFQAGVSMGCGAORNRPGVYTRLPFLFRDWIKENTG 240
OY	795	SGVDSCGSGGPGPLSSVEADGRIFQAGVSMGCGAORNRPGVYTRLPFLFRDWIKENTG 854
DB	241	V 241

QY 855 V 855

RESULT 2

ID R57283 standard; Protein: 798 AA.

AC R57283;

DT 08-MAR-1995 (first entry)

DE Bovine enterokinase.

KW Enterokinase; EK; heavy chain; light chain; catalytic domain; digestive disorder; cleavage; fusion protein; trypsinogen; trypsin; enzyme; PACE gene.

OS Bos taurus.

PI Key

PI domain

PI location/Qualifiers

PI 1..563

PI /label= "heavy-chain-C-terminal"

PI /note= "non-catalytic domain"

PI 564..798

PI /label= "light-chain"

PI /note= "catalytic domain"

PN W09416083-A.

PN 21-JUL-1994.

PN 15-JAN-1994; 000616.

PN 15-JAN-1993; US-005944.

PA (GEMT) GENETICS INST INC.

PI Lavalie ER;

PI WPI: 94-249229/30.

DR N-PSDB: Q70104.

PT New nucleic acid encoding enterokinase activity - and related vectors, host cells, expression products and antibodies are useful in treating digestive disorders and for cleaving fusion proteins

PT Proteins

PS Disclosure; Page 28-30; 50pp; English.

CC The enterokinase (EK) (or the EK gene when used in gene therapy) is used to treat digestive disorders associated with low EK activity (esp. inability to process trypsinogen to trypsin). For cleaving fusion proteins, recombinant EK catalytic domain is much more efficient than the native two-chain holoenzyme and is not contaminated by other proteolytic enzymes. For expression of recombinant EK, the 1691-2398 DNA fragment was fused to the 3'-end of the signal peptide and pro-region of the human PACE gene. The prod. could be expressed in CHO cells to produce a chimeric prod. from which the pro-region as cleaved by endogenous PACE, providing mature EK catalytic domain.

CC which the pro-region as cleaved by endogenous PACE, providing mature EK catalytic domain.

CC Sequence 798 AA;

QY Query Match

QY Best Local Similarity 42.9%; Score 768; DB 1; Length 798;

QY Matches 102; Conservative 61; Mismatches 68; Indels 7; Gaps 6;

QY 562 PKIVGSDSREGAMPVVALY-FDDQVCASIVSRDMLVSAHC-VY--G-RNNEPSKV 616

QY 613 ARVVGSTDADEGEMPMQVSLHALGCHIGASISPMVLVSAHCYIDDRGRYSPTQW 672

QY 617 KAVLVGMAASNLTPQIEPRLLDQIVINPHYKRRKRNNDIAMHLEMKYNTDYDIPCL 676

QY 673 TAVLGHDOSORSAFGVQSRRLKRIISHPFDFDIDALELEKPEYSSMVRPDL 732

QY 677 PEENOVFPGRICISAGMGALTYOGSTADVLQADVPALISNEMCOQOQMEVNTENMCA 736

QY 733 PDSHFRPGKAIKWTGKHGTQYGGALILQKGEIRVINTTCENLQO-QITPRMKCV 791

QY 737 GTEAGVSDSCQDGGPPLMCOENRML-AGVTSFGYCALPRRPGVYARVPRTEWI 793

QY 792 GFLSGVDSCQDGGPPLSSVEADGRIFGAGVYVSMGDCGCAQRNKPQVYRLPLFRDWI 849

RESULT 3

ID R89435 standard; Protein: 418 AA.

AC R89435;

DT 27-MAY-1996 (first entry)

DE Trypsin-like enzyme.

KW Trypsin; fibrinogen; thrombin; expectorant; respiratory disease; asthma; VIP; vasoactive intestinal peptide; Influenza virus;

KW protease; primer; PCR; amplification.

OS Homo sapiens.

PI Key

PI peptide

PI 1..186

PI /label= "sig-peptide"

PI 187..418

PI /label= "mat-protein"

PN AU9527248-A.

PN 08-FEB-1996.

PN 31-JUL-1995; 027248.

PN 29-JUL-1994; JP-178607.

PA (TEIJ) TEIJIN LTD.

PI Masuda K, Ogawa H, Suga T, Sugimoto Y, Takagi K;

PI WPI: 96-117356/13.

DR N-PSDB: T10689.

PT Nucleic acid sequence encoding trypsin-like enzyme - which digests fibrinogen, used as expectorant in treatment of respiratory diseases, e.g. bronchial asthma

PS Example 11; Page 47-49; 65pp; English.

CC The overlap parts of 107 bp between the sequences given in T10698 and T10703 were identified, and thereby their identity was confirmed. It was confirmed from the sequence analysis, that these overlapping sequences contained a region encoding the amino acids of the N-terminus 20 residues of the trypsin-like enzyme isolated from the cough phlegm. The sequences were ligated, and the desired trypsin-like enzyme gene cDNA sequence was determined (T10689).

CC Sequence 418 AA;

QY Query Match

QY Best Local Similarity 40.0%; Score 677; DB 1; Length 418;

QY Matches 104; Conservative 64; Mismatches 76; Indels 16; Gaps 12;

QY 111 NECGAPDLITSEPRILGTEAEGRSPQVSLR-LNNAHCGSLNNMMILTAACF 229

QY 600 KDCDCGLSFT-ROARVVGSTDADEGEMPMQVSLHALGCHIGASISPMVLVSAHCY 658

QY 230 ---RS-N-SNPDVATSGI---SN-TFPKR-MRVRNLIHNKSKATHEIDALVRL 279

QY 659 IDDRFRTSDPTQWTAFLGLHDOSORSAFGVQSRRLKRIISHPFDFDIDALELE 718

QY 719 KPAEYSMVRPCLPDASHVPAKAIWGTGKHGTQYGGALILQKGEIRVINTTCEN 778

QY 340 PHSYVGLSLGMLCAGVQGGVDACQDGGPPLVQ-EDSRMLRIVGVSMGDCGLPDK 398

QY 779 LFP-Q-QITPRMKCVGFLSGVDSCQDGGPPLSSVEADGRI-FGAGVYVSMGDCGCAQRN 835

QY 399 PGVYRTAVLDMIRQNGI 418

QY 836 PGVYRLPLFRDWIENKGV 855

RESULT 4

ID R89430 standard; Protein: 232 AA.

AC R89430;

DT 26-MAY-1996 (first entry)

DE Trypsin-like enzyme.

KW Trypsin; fibrinogen; thrombin; expectorant; respiratory disease; asthma; VIP; vasoactive intestinal peptide; Influenza virus;

OS Homo sapiens.

PI Key

PI peptide

PI 1..20

PI /label= "N-terminal"

PN AU9527248-A.

PN 08-FEB-1996.

PN 31-JUL-1995; 027248.

PN 29-JUL-1994; JP-178607.

PA (TEIJ) TEIJIN LTD.

PI Masuda K, Ogawa H, Suga T, Sugimoto Y, Takagi K;

PI Yamaoka K, Yasuoka S;
 DR WPI: 96-117356/13.
 DR N-PSDB: 110688.
 PT Nucleic acid sequence encoding trypsin-like enzyme - which digests fibrinogen, used as expectorant in treatment of respiratory diseases, e.g. bronchial asthma
 PS Claim 2; Page 56; 65pp; English.
 CC The nucleic acid may be used in the prodn. of the trypsin-like enzyme. The enzyme digests synthetic trypsin and thrombin substrates, fibrinogen (therefore used as an expectorant in the treatment of respiratory diseases, eg bronchial asthma) and vasoactive intestinal peptide, but does not digest IgA, IgG, albumin, alpha-1-antitrypsin or substance P.
 CC In construct to trypsin, the enzyme also inactivates influenza viruses, the Miyadera strain of NDV and the New Jersey strain of VZV.
 CC Sequence 232 AA;
 Query Match 10.4%; Score 663; DB 1; Length 232;
 Local Similarity 41.0%; Pred. No. 3,01e-49;
 hes 100; Conservative 60; Mismatches 69; Indels 15; Gaps 11;

DB 1 IIGGEAEFGSPWQVSLR-LNNAHGCGSLNNMMILTAHCF---RS-N-SNPRDWA-54
 QY 615 VVGTDADGEMPMQVSLHALGQGHICGASLSPMVLAAHCYIDRFRISDPTQWTA 674
 DB 55 TSGI--ST-TPEKLR-MKRNILHNKYSATHEMDIALVRLNSVFTKDIHSCLPA 109
 QY 675 FTGLHDGOSRSPGVQERLKRILSHPFNDFTFYDIALLEKPAEYSSNVRICLPD 734
 DB 110 ATQNTIPGSTAVYTGMAOEYVGHVPELRQGVRIISNDVCAHNSGAILSLCAG 169
 QY 735 ASHVFPAGKAIWVTGSHGYGTGALLQKGEIRINOTTCENLFP-Q-QITPRMCGV 792
 DB 170 VPGGVADCGSGGPELVQ-EDSRRLMEIVGIVSGDOGLDKRQVYTRVAYLDWIRQ 228
 QY 793 FTSGGVDSQSGSGPLSSVEADGRI-FQAGVSWGDCAGKQKRPVTRPLERDWIKE 851

DB 229 QTGI 232
 QY 852 NTGV 855

RESULT 5
 ID W96812 standard; Protein: 416 AA.
 AC W96812;
 DT 21-APR-1999 (first entry)
 DF A mouse serine protease called hepsin.
 KW mouse serine protease; hepsin; animal model; bone disease; bone disorder; skeletal disorder; osteoporosis; Paget's disease; osteitis deformans; elevated bone alkaline phosphatase level.
 OS Mus musculus.
 PN MO9854307-AL.
 PD 03-DEC-1998.
 PF 29-MAY-1998; E03199.
 PR 30-DEC-1997; US-000486.
 PR 30-MAY-1997; US-866058.
 PA (SCHD) SCHERING AG.
 PA (UNIM) UNIV WASHINGTON.
 PI Sadler JE, Wu Q;
 DR N-PSDB: X15134.
 DR WPI: 99-070213/06.
 PT New nucleic acid functionally disrupts mouse hepsin gene - used to provide transgenic mice with abnormally elevated blood alkaline phosphatase, useful as models for bone disorders
 PS Example 1; Fig 2; 29pp; English.
 CC The present sequence represents a mouse serine protease called hepsin. The specification describes a mammalian cell in which expression of a gene encoding hepsin has been functionally interrupted or suppressed. The products and methods provide an animal model for bone disease, and are useful to determine effective treatment for bone and skeletal disorders such as osteoporosis, Paget's disease and osteitis deformans, especially those associated with elevated bone alkaline

CC phosphatase levels.
 SQ Sequence 416 AA;
 Query Match 9.7%; Score 618; DB 1; Length 416;
 Best Local Similarity 39.5%; Pred. No. 5.72e-45;
 Matches 102; Conservative 55; Mismatches 84; Indels 17; Gaps 16;

DB 149 CQDCGRRLPYD-RIVGQDSSLGRMPQVSLRYDG-THLGSSLLSGDWLTAHCF-P 205
 QY 602 C-DGLRSFTVQARVGVGTDADEGEMPMQVSLHALGQGHICGASLSPMVLAAHCYID 660
 DB 206 ERN-R-V-LGMRVFPAGVARTSPHAYOLGYQAVTY-HGGLPRPDRPTIDNSNDIALYH 261
 QY 661 DRGFRYSPTQWTFELGHDQ-SQSRAP-GVQERRLRKIIISHPFN-DFTD-YDIALE 716
 DB 262 LSSSLPLREXIOPLCPAAGALVDGKVCVTGNGNTQFYGOQAMVLOEAPVSIISNEVC 321
 QY 717 LEKPAEYSSNVRPFCICLPASHVFPAGKAIWVTGSHGYGTGALLILOKGEIRINOTTC 776
 DB 322 NSPDPYGNQIRPKMFCAGYRPGGIDACGDSGPFVCEDSISGRWRLGIVSWGTCGA 381
 QY 777 E-NLLPQQTTPRMVCVGLSGVDSCGSGPLSSVEA-DG-RIFQA-GVSWGDCGA 831

DB 382 LARKPGVTTATYDREMT 389
 QY 832 QNRKPGVTTATYDREMT 849

RESULT 6
 ID W46917 standard; Peptide: 356 AA.
 AC W46917;
 DT 02-JUL-1998 (first entry)
 DE Amino acid sequence of a novel human kallikrein.
 KW Kallikrein; HKLP; human; serine protease; drug screening; agonist; treatment; hypertension; cardiac hypertrophy; arthritis; inflammatory disorder; blood clotting disorder.
 KM Homo sapiens.
 OS Homo sapiens.
 FH Key
 FT Misc_difference 106
 FT Location/Qualifiers
 FT Misc_difference 168
 FT /note- "encoded by YGC"
 FT /note- "encoded by YGC"
 FT W09803665-AL.
 PD 29-JAN-1998.
 PF 21-JUL-1997; U.2724.
 PR 22-JUL-1996; US-681151.
 PA (IMCY-) INCYTE PHARM INC.
 PI Au-Young J, Bandman O, Braxton SM, Goll SK;
 DR WPI: 98-120785/11.
 PT Human kallikrein polypeptide and DNA encoding it - useful for screening compounds useful in treatment of e.g. hypertension, cardiac hypertrophy and arthritis
 PS Claim 1; Fig 1A-E; 59pp; English.
 CC The present sequence represents a novel human kallikrein (HKLP). Kallikreins are a large family of homologous serine proteases that act in a variety of circulatory and immune system functions. The cDNA sequence encoding HKLP was isolated from a heart tissue library, and was identified in cDNA. Incyte clone 307474, through a computer generated search for amino acid sequence alignments. HKLP shows 35% identity to rat kallikrein and 33% identity to human kallikrein. Unlike rat kallikrein, HKLP is hydrophobic in the carboxy terminus and likely to remain membrane bound. The HKLP protein has 3 potential glycosylation sites. The cDNA sequence encoding HKLP, and vectors and host cells containing it are useful for the recombinant production of HKLP. HKLP is useful in drug screening for potential antagonists (or agonists). The HKLP protein, cDNA and cDNA are useful in the treatment of conditions such as hypertension, cardiac hypertrophy, arthritis, inflammatory disorders and blood clotting disorders.
 CC Sequence 356 AA;
 Query Match 9.6%; Score 615; DB 1; Length 356;
 Best Local Similarity 37.3%; Pred. No. 1.10e-44;

FT	Disulfide_bond	202	/note- likely to be involved in disulphide bonding"
----	----------------	-----	---

DR WPI: 98-393476/34.
PT Human plasminogen derived polypeptide - has neovascularisation
PT inhibiting activity
PS Claim 1: Page 2; 16pp: Japanese
CC The invention relates to a neovascularisation inhibitor which comprises
CC amino acids 355-791 of human plasminogen. Also claimed are a method for
CC the preparation of angiotatin, and angiotatin prepared by this method.
CC The human plasminogen protein fragment is prepared by: (a) applying human
CC plasminogen to a lysine sepharose column to separate it into plasminogen
CC form 1 and form 2; (b) separating plasminogen form 1 and form 2 and
CC digesting them with elastase; (c) fractionating the elastase-decomposed
CC product of form 1 plasminogen and form 2 plasminogen in a lysine
CC sepharose column; (d) collecting the fractions bound to the lysine
CC sepharose column; (e) further fractionating the form 2 plasminogen using
CC an aminoheaxal sepharose column; and (f) collecting the fraction bound to
CC the aminoheaxal sepharose column. This human plasminogen fragment can be
CC used to inhibit growth of vascular endothelial cells. The present
CC sequence represents amino acids 355-791 of human plasminogen.
SQ Sequence 437 AA;

ry Match 8.9%; Score 566; DB 1; Length 437;
c Local Similarity 43.3%; Pred. No. 4.79e-40;
Matches 104; conservative 47; Mismatches 71; Indels 18; Gaps 13;

Dd 206 GRVVGCGVAHPHSWQVSLRTRFGC-HFCGGTILSPENVLTAAHC-LE-KSPR---PSS 259
::: | | | | | : | | | | | : | | | | | : | |
QY 613 ARVVGGTDADGEMWQVLSLHA-LGGCHICGASLISPMVLSAAHCYIDRGFRYSDPQT 671

```

Db      260 YKVIILGAHQEVNL-EPHVQE--IE-VSRLELEP-TRK-DIALIKLSSPAVITDKYIPAC 312
      :||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Oy      672 WTAFELGLHQSQSRSAFGVGERLKRITLSHPFNDFTFDYDIALLELEKPRAYSSSWYRPIIC 731

```

D0 313 LPSBNVAVADRTECITGMGTQ-GTFGAGLLKEAQLPVIENRYCNREYELNGRVOSTEL 371
 | : : : : ||| | | | : : : : :
OY 732 LPDASHVFPAKAIWTGWHTGYGGCGALLDKGEIRINOTICE--NLPPQTTPRRM 789

Dd 372 CAGHLAGGIDSCQGDSGGPLVCFEKRYILO-GVTSMGLGCARPKNPGYIVRVSRPWTWI 430
 | : | : | | | | | | | : : | | | | : |
Oy 790 CVGFLLSGVDSCQGDSGGPLSSVEADGRIFQAGVSWMGCGCAQNKPgyITRLPLERDWtI 849

RESULT	14
ID	R34427 standard; Protein; 546 AA.

DT 17-AUG-1993 (first entry)
DE Sequence of tissue plasminogen activator (t-PA)/plasminogen
DE hybrid protein.

0 Synthetic.
I J55200340-A.
PL 06-APR-1993.

PR 22-MAY-1987; US-053412.
PA (ZYMO) ZYMOGENETICS INC.
PI Foster DC, Mulvihill ER, Ohara PJ, Pingel K, Yoshitake S;

DR	N-PSDB; Q40318,
PT	Human tissue plasminogen activator single chain form fibrinolytic
PT	agent - compiles thrombin cleavable zymogen stimulating amido

PT victims and suppressing fibrin matrix
PS Example: Fig 8A, 8B, 8C, 22pp; English.
CC A hybrid DNA sequence was constructed which encoded a protein

the cysteine in posn. 261) joined to the serine protease domain of plasminogen beginning at amino acid 541 (just to the amino-terminal side of the normal activation site). This hybrid protein

SO Sequence 546 AA; 8.98; Score 566; DB 1; Length 546; Query Match

Matches 104; Conservative 47; Mismatches 71; Indels 18; Gaps :

```

Dd      315 GRVYGGCAAPHSMNPMOVSLETRGCM-HFGQGTISEWUUTLAHC-LE-KSPF---PSS    368
           ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
Oy      613 ARVVGDTDAEGCEPNMVOYSLHA-LGQHICGASLSISRWLVSAHACITDDRGFFYSPTQ   671
Dd      369 KYVLIGAHOEVLN-EPHYOE--IE-VSRLEED-TRK-DIALKLKSSPAVITDKVIPAC   422
           :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :
Oy      672 WTAEFGIHDGQSRSAPGOERRLRRIISHPEFNDFDYDIALLEKEPPAYESSMVAPIC     731
Dd      422 LPSRYVVADTEEFETINGMBETO-CTIGAGLLKKAOLPVLENKYCNXREFLNGVOSTEL   480
           ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :
Oy      732 LPDSHFHPACKALMTVWGHWGTQTGGTGALLQGELRVINQTTCE--NLIPQDITPRM     789
Dd      481 CAGHLAAGTDSGCQSDSGGLVCFEFKDXLYILO-GVTSMGLCAREPKNEGVVRVRSFTWI     539
           ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :
Oy      790 CVGFSLSGVDSCQGDSDSGPLSVSABDRIRPGAGVYVMGDCQAQNKKRGVYTRLPFLFDWI   849

```

RESULT	15
ID	R60519 standard; Protein; 790 AA

DT	22-MAR-1995 (first entry)
DE	Human 'Glu' plasminogen.
KW	Serine protease; Factor-Xa; recognition site; plasminogen; kringles

KW polymerase chain reaction; amplification.
 OS Homo sapiens.
 PN W09418227-A.
 RD 18-MTC-1004

PF 04-FEB-1994; DK0054.
PR 04-FEB-1993; DK-000130.
PR 05-FEB-1993; DK-000139.
03 FEB 1993 DK 000133
03

PA (DENE-) DENEZYME APS.
PI Etzerodt M, Holtet TL, Thøgersen HC;
DR WPI: 94-279681/34.
Dose of 94-279681/34.

PT involving denaturing and renaturing conditions to produce a
PT correctly folded prod
PS Disclosure; Page 148-50; 202pp; English.

CC sequence given in R60519) was PCR amplified using primers given in
CC Q71268-71. Amplified cDNA was linked to a sequence encoding the
CC Factor-Xa cleavage site (given in R60503), subcloned in vector

CC sequence and expressed in *E. coli* QY13. The fusion protein was
CC purified on an N12+-activated NTA-agarose column. A cyclic
CC procedure was used to obtain correctly folded recombinant protein.

Query Match 8.98; Score 566; DB 1; Length 790;
Best Local Similarity 43.38; Pred. No. 4.79e-40;
Fast Local Similarity 42.38; Pred. No. 1.1e-38

```
Db      559 GRVGGCAHPRHPWQYSLRFRGM-HECGGLISPELVLTAAHC-LE-KSPR---PSS 6
```

Db 613 YKVILGAHQEVL-EPHYQ--IE--VSRLFLER-TRK-DIALKLSSPAVITDKVIPAC 6

Db 666 LPSPIYVADRIECFITGGETQ-GTFGAGILKEAQLPIENKVCNRYEFLNGRQSTEL 7

D6 725 CAGHLAGTDSQGDSDGLVCFEKKYILQ-GVTSWGLGACARPKNKPGVYVVRVSRFVTNI 7

Search completed: Thu Apr 27 09:33:29 2000
Job time : 87 secs.

RESULT 15
ID R60519 standard; Protein; 790 AA.
AC R60519; 22-MAR-1995 (first entry)
DT Human 'Glu' plasminogen.
DE Serine proteases; Factor-Xa; recognition site; plasminogen; kringle;
KW fusion protein cleavage; protein folding; primer;
KM polymerase chain reaction; amplification.
OS Homo sapiens.
PN W09418227-A.
PD 18-AUG-1994.
PF 04-FEB-1994; DK0054.
PR 04-FEB-1993; DK-000130.
PR 05-FEB-1993; DK-000139.
PR 03-DEC-1993; WO-G02492.
PA (DENZ-) DENZYME APS.
PI Etzerodt M, Hollet TL, Thøgersen HC;
DR WP1; 94-279681/34.
PT Refolding of polypeptide molecules - using a cyclic process
PT involving denaturing and renaturing conditions to produce a
PT correctly folded prod
PS Disclosure; Page 148-50; 202pp; English.
CC cDNA encoding kringle domains 1 and 4 of human plasminogen (full
CC sequence given in R60519) was PCR amplified using primers given in
CC Q71268-71. Amplified cDNA was linked to a sequence encoding the
CC factor-Xa cleavage site (given in R60503), subcloned in vector
CC pUC19/M6 so that it was linked to a hexahistidine-encoding
CC sequence and expressed in *E. coli* O13. The fusion protein was
CC purified on an Ni2+-activated WTA-agarose column. A cyclic
CC procedure was used to obtain correctly folded recombinant protein.
CC Sequence 790 AA;

Query Match	8.98;	Score 566;	DB 1;	Length 790;
Best Local Similarity	43.38;	Pred. No. 4.79e-40;		
Matches	104;	Mismatches 71;	Indels 18;	Gaps 13
	Conservative			

Db 559 GRVGGCVAAHPHSPWQVSLRTREGN-HFCGGLISPEWLTAANC-LE-KSPR--PSS
 :||||| | |||||::: | :||::|||:||:||||: |:: |
Qy 613 ARVVGTDALDEGEPMQVSLHA-LQGCHICGASLISPNNLVSAHCYIDDRGFRRSDPTQ

D6 613 YKVIIGSAHOEVLN-EPHVOE--IE--VSRLTEP-TRK-DIALKISSPAVITDKVIPAC 6
 :||:-:| ||: |::|::| ||||| ||:|:| |
 QY 672 WTAIFGLHDQSORSAPGVQGERLKRITISHPFNDFTFDIDALLLEKEPAEYSMMVRPIC 7

D6 666 LPSPNVADRTCECFITGNGETO-GTGGAGLKEAQLPIENKVCNRYEFLNRQVSTEL 7
||::| : |||| | | | : :: || : : : :
QY 732 LPDASHVEFPAGKAIVWTGHTGYGTGALILOKGEIRINOTTC--NLPPQIITPRM 7

D_b 725 CAGHLAGGTDSCGGDSCGELVCFEEXDKYILQ-GVTSMGLGARPNKPGVYYVRVSFRFTWI 7
| | : | | | | | | | | : | | | | | | | : | |
QY 790 CVGFSLSGVDSCQGDSGFLSVSEADGRIFPAQGVVSWMGDCAQRRNKPVGYYTRPLPFRDWI 8

Search completed: Thu Apr 27 09:33:29 2000
Job time : 87 secs.

45 558

Plasminogen mutein T8 2.7.2000

ALIGNMENTS

RESULT 1
ID W22987 standard; Protein; 241 AA.
AC W22987;
DT 08-OCT-1997 (first entry)
DE Human serine protease 67 (SP67).
KW Human; colon carcinoma; COLO 201; cell line; serine protease; SP67;
screening; inhibitor; treatment; disease.
OS Homo sapiens.
PN J09149790-A.
PD 10-JUN-1997.
PF 24-JUL-1996; 212196.
PR 29-SEP-1995; JP-275105.
PA (SUNR) SUNTORY LTD.
DR WPI; 97-357902/33.
DR N-PSDB; T79128.
PT Human colon carcinoma derived serine protease(s) SP59, SP60 and SP67
PT - useful to screen for specific inhibitors, e.g. to search for, or
PT study agent for treatment of various diseases
PS Claim 1; Pages 12-13; 16pp; Japanese.
CC The present sequence is the human colon carcinoma COLO 201
CC cell line derived serine protease 67 (SP67), which can be used to
CC screen for specific inhibitors, e.g. to search for, or study an
CC agent for the treatment of various diseases.
SQ Sequence 241 AA;

Query Match 28.3%; Score 1808; DB 1; Length 241;
Best Local Similarity 99.6%; Pred. No. 8.25e-161;
Matches 240; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 VVGGTDADEGEWPQVSLHALGQGHICGASLISPNWLVSAAHCYIDDRGFRYSIPTQWTV 60
QY 615 VVGGTDADEGEWPQVSLHALGQGHICGASLISPNWLVSAAHCYIDDPGPPYSIPTQWTA 674
Db 61 FLGLHDQSQRSA PGVOERRLKRIISHPFFNDFTFDYDIALLELEKPAEYSSMVPPICLPD 120
QY 675 FLGLHDQSQRSA PGVOERRLKRIISHPFFNDFTFDYDIALLELEKPAEYSSMVRPICLPD 734
Db 121 ASHVFPAGKAIWVTGWGHTOYGGTGALILQKGEIRVINOTTCENLLPQOITPRMVCVGF 180
QY 735 ASHVFPAGKAIWVTGWGHTOYGGTGALILQKGEIRVINOTTCENLLPQOITPRMVCVGF 794
Db 181 SGGVDSCQDSSGGLSSVEADGRIFQAGVVSNGDGCAQRNKPQVYTRPLPLFRDWIKENTG 240
QY 795 SGGVDSCQDSSGGLSSVEADGRIFQAGVVSNGDGCAQRNKPQVYTRPLPLFRDWIKENTG 854
Db 241 V 241